

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2001-309225

(43)Date of publication of application : 02.11.2001

(51)Int.Cl.

H04N 5/225
G03B 17/24
G03B 17/48
G03B 19/02
G06T 1/00
G06T 7/00
H04N 5/907
// H04N101:00

(21)Application number : 2001-046640

(71)Applicant : EASTMAN KODAK CO

(22)Date of filing : 22.02.2001

(72)Inventor : RAY LAWRENCE A
NICPONSKI HENRY

(30)Priority

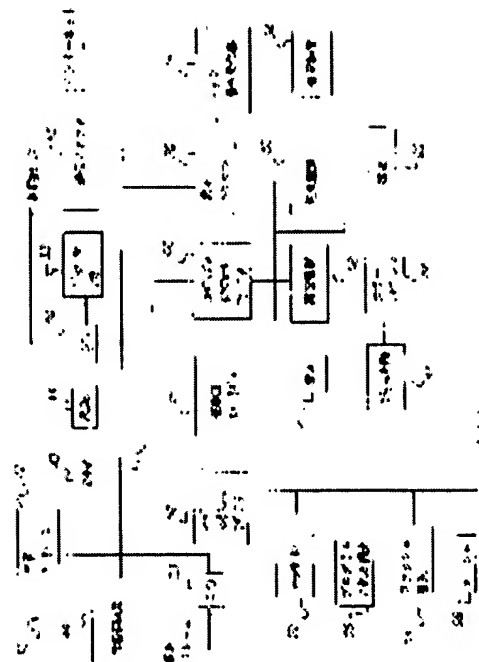
Priority number : 2000 514436 Priority date : 28.02.2000 Priority country : US

(54) CAMERA FOR DETECTING FACE AND ITS METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To detect more than one face in an image so as to improve the quality of a captured image by adding a value to image capturing processing.

SOLUTION: A method for determining the presence of a face from image data is performed by comparing facial models for searching more than one face candidate regions with facial probabilities to previously screen the image data with a first algorithm, and by performing operation on the face candidate region with a second algorithm by pattern matching so as to check each face candidate region and to confirm the presence of a face in the region. As for detection level, a combination of the first and second algorithm has a higher performance than the individual algorithm. A digital camera includes a memory for storing an algorithm and an electronic processing unit for processing image data with algorithm so as to determine the presence of more than one face in the scene. The facial data is used for controlling exposure parameter of image capturing or for producing processed image data that associates color balance to the presence of faces. Alternatively, it can be stored on a storage medium together with the image data.



LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

【特許請求の範囲】

【請求項1】 シーンの画像を捕捉するデジタルカメラであって、
画像を捕捉し、画像データを生成する捕捉部と、
上記シーン中の1つ以上の顔の存在について判定するために上記画像データを処理する電子処理部と、
上記処理部に関連付けられ、上記画像中の少なくとも1つの顔の属性に対応する顔データを発生する顔データ手段と、
上記画像データを記憶する記憶媒体と、
上記処理部に関連付けられ、上記顔データを上記画像データと共に上記記憶媒体上に記録するための記録手段とを含むデジタルカメラ。

【請求項2】 上記顔データは、上記画像中の少なくとも1つの顔についての位置、向き、スケール、及びポーズのうちの少なくとも1つに対応する、請求項1記載のデジタルカメラ。

【請求項3】 電子媒体と磁性層を有するフィルム媒体の両方にシーンの画像を捕捉するハイブリッドカメラであって、
画像センサによって画像を捕捉し画像データを生成する画像捕捉部と、
上記フィルム媒体上に上記画像を捕捉する手段と、
上記シーン中の1つ以上の顔の存在について判定するために上記画像データを処理する電子処理部と、
上記電子処理部に関連付けられ、上記画像中の少なくとも1つの顔の位置、スケール、及びポーズのうちの少なくとも1つに対応する顔データを発生する顔データ手段と、
上記フィルム媒体の磁性層上に上記顔データを書き込む手段とを含むハイブリッドカメラ。

【請求項4】 上記画像データを記憶する記憶媒体と、
上記処理部に関連付けられ、上記顔データを上記画像データと共に上記記憶媒体に記録する記録手段とを含む、請求項3記載のハイブリッドカメラ。

【請求項5】 画像データから顔の存在について判定する方法であって、

(a) 画像の1つ以上の顔候補領域を見つけるために、顔形状モデルと上記領域中の画素に割り当てられた顔確率との比較に基づいて、第1のアルゴリズムにより画像データを予めスクリーニングする段階と、

(b) 上記画像の各顔候補領域を表わし、上記領域中の顔の存在を確かめるために、パターンマッチング技術を用いて、第2のアルゴリズムにより顔候補領域に対して演算を行なう段階とを含み、
従って検出レベルに関しては、上記第1及び第2のアルゴリズムの組合せは、いずれかのアルゴリズムを個々に用いた場合よりも高い性能を与える方法。

【請求項6】 画像中の赤目の存在を確認する方法であって、

画像を表わす画像データを与える段階と、

上記画像中に赤目が存在することを示す赤目信号を発生するために赤目検出アルゴリズムにより上記画像データを処理する段階と、

上記赤目信号が上記画像中に1つ以上の顔が存在することに対応することを確認するために顔検出アルゴリズムを用いることによって赤目の存在を確証する段階とを含む方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は画像捕捉の分野に関連し、特にカメラによる画像捕捉の処理を改善し最適化するための画像処理分野に関する。

【0002】

【従来の技術】撮影者によって収集される大多数の画像は、多くの場合に画像の最も重要な被写体である人物を含む。画像中での人物の存在及び位置、特に人物の顔の存在及び位置を知ることは、画像捕捉処理において多くの有益な改善がなされることを可能とする。その幾つかについては従来技術で提案されている。例えば、自動式及び半自動式合焦カメラは、最良焦点として調整されるシーンの部分を選ぶことが多い。カメラがシーン中で顔の位置を見つけることができれば、撮影者がその選択をはっきりと無効としないかぎり、焦点は顔について最適化される。米国特許第5,835,616号では、顔検出システムは、自動写真技術において、合焦された被写体がないために低い質を生じさせる手動調整問題を除去するために使用される。

【0003】更に、シーン中の顔の検出は、主要な被写体の正しい位置についての非常に強い証拠を与える。これに関して、第616号特許に開示される処理は、デジタルカメラによって撮影されたデジタル化された画像中の人間の顔を自動的に見つけ、顔特徴について調べることによって顔の存在を確かめ、カメラは検出された顔に中心が合わされる。顔の検出は、また顔及び／又は肌の領域について正しいカラーバランスの強い証拠を生じさせる。例えば、米国特許第5,430,809号では、ビデオカメラは、自動露出及び自動焦点のために顔対象に測定フレームを設定するため顔目標を自律的に追跡する。更に、一旦測定フレームが設定されると、顔の最適な肌の色を得るために自動白バランスシステムが色を調整する。結果として、自動白バランスシステムは、自動肌カラーバランスを行なうものであるといえる。また(米国特許第5,629,752号より)、人間の顔を検出し、次に顔に対応する領域が写真印刷装置によって適切に印刷されよう、露出量を決定するために顔領域の色及び／又は密度を表わすデータを使用することが知られている。

【0004】顔検出は、過去数年に亘り画像理解の主題に関連して研究されてきたが、特に強固な顔検出アルゴ

リズムが必要とされる場合はかなりの計算が必要となる場合がある。様々な撮像条件に亘ってかなり良い性能を示す多数の方法が考えられてきた。かかる方法は、(ハンドヘルドカメラと比較して)比較的高度な処理能力を有する写真プリンタといった大規模処理機器においてよりうまく実施されうる。これらの顔検出方法を、限られたメモリ資源を有するカメラで、低い計算費用で良く実施することが困難であった。これがうまく行なわれれば、シーン中での顔の検出は、画像捕捉処理における多数の改善へのきっかけとなるであろう。更に、画像捕捉の後の下流の活動を実施するために顔を検出することが有用であり、例えば顔検出は続く印刷(例えばインデックスプリント)についての上下方向の証拠を与えうる。

【0005】

【発明が解決しようとする課題】本発明は、画像を捕捉し、画像捕捉処理に対して価値を付加すると共に捕捉された画像の質を改善するために画像に含まれる1つ以上の人間の顔を検出することを目的とする。本発明は、上述の1つ以上の課題を解決することを目的とする。

【0006】

【課題を解決するための手段】概説するに、本発明の1つの面によれば、画像データから顔の存在について判定する方法は、2つの別々のアルゴリズム段階を有する顔検出アルゴリズムを含み、即ち、第1の段階では画像の1つ以上の顔候補領域を見つけるために、顔形状モデルと上記領域中の画素に割り当てられた顔確率との比較に基づいて、第1のアルゴリズムにより画像データを予めスクリーニングし、第2の段階では、上記画像の各顔候補領域を表わし、上記領域中の顔の存在を確かめるために、パターンマッチング技術を用いて、第2のアルゴリズムにより顔候補領域に対して演算を行ない、従って検出レベルに関しては、上記第1及び第2のアルゴリズムの組合せは、いずれかのアルゴリズムを個々に用いた場合よりも高い性能を与える。

【0007】カメラとしての実施では、デジタルカメラは上述の第1及び第2のコンポーネントを含むアルゴリズムを記憶するアルゴリズムメモリと、シーン中の1つ以上の顔の存在について判定するためにアルゴリズムと共に画像データを処理する電子処理部とを含む。顔の存在を示す顔データは、例えば画像の捕捉の露出パラメータを制御するため、又は例えばカラーバランスを画像中の顔の存在に関連付ける処理された画像データを生成するために使用されてもよく、又は顔データは記憶媒体上に画像データと共に記憶されうる。

【0008】本発明の他の面では、デジタルカメラは、画像を捕捉し、画像データを生成する捕捉部と、上記シーン中の1つ以上の顔の存在について判定するために上記画像データを処理する電子処理部と、上記処理部に関連付けられ、上記画像中の少なくとも1つの顔の属性に対応する顔データを発生する顔データ手段と、上記

画像データを記憶する記憶媒体と、上記処理部に関連付けられ、上記顔データを上記画像データと共に上記記憶媒体上に記録するための記録手段とを含む。かかる顔データは、上記画像中の少なくとも1つの顔についての位置、向き、スケール、又はポーズに対応する。

【0009】本発明の更なる面では、デジタルカメラは、画像データ中の1つ以上の顔の存在について判定する顔検出アルゴリズム及び幾つかの所定の構図原則に基づく構図調整を提案する構図アルゴリズムと、シーン中の1つ以上の顔の存在と、及びそれらと幾つかの所定の構図原則との関係とを決定するアルゴリズムと共に画像データを処理する電子処理部とを含む。次に処理部は、画像中の少なくとも1つの顔の位置、向き、スケール又はポーズに対応する顔データと、所定の構図原則からの顔データのずれに対応する構図案とを発生する。

【0010】本発明の更なる面では、電子媒体と磁性層を有するフィルム媒体の両方にシーンの画像を捕捉するためのハイブリッドカメラが開示される。ハイブリッドカメラは、電子媒体と磁性層を有するフィルム媒体の両方にシーンの画像を捕捉するハイブリッドカメラが開示される。ハイブリッドカメラは、画像センサによって画像を捕捉し画像データを生成する画像捕捉部と、上記フィルム媒体上に上記画像を捕捉する手段と、上記シーン中の1つ以上の顔の存在について判定するために上記画像データを処理する電子処理部と、上記電子処理部に関連付けられ、上記画像中の少なくとも1つの顔の位置、スケール、及びポーズのうちの少なくとも1つに対応する顔データを発生する顔データ手段と、上記フィルム媒体の磁性層上に上記顔データを書き込む手段とを含む。

【0011】本発明の更なる面では、本発明は、デジタル画像センサ、マイクロプロセッサといった中央処理装置、顔を検出する手段、及び検出された顔の位置を撮影者に対して表示する手段を組み込んだカメラを含む。

【0012】

【発明の実施の形態】本発明の上述及び他の目的、面、特徴及び利点は、添付の図面を参照して以下の望ましい実施例についての説明及び請求項により、より明らかに理解され認識されるであろう。

【0013】電子及びフィルム捕捉を使用した撮像システムは周知であるため、本願の記載は、特に、本発明によるシステム及び装置の部分をなす、又はより直接的に関連する属性について説明するものとする。本願に特に図示されず記載されないシステムの属性及び構成装置は、技術分野において知られているものから選択されうる。以下の説明では、顔検出アルゴリズムの望ましい実施例は、通常はソフトウェアプログラムとして実施されうるが、当業者はかかるソフトウェアと同等のものがハードウェア中に構築されうることを容易に認識するであろう。以下記載されるシステム及び方法が与えられれば、本発明を実施するのに必要とされる全てのかかるソ

フトウェア実施は従来通りであり、かかる技術分野における通常の技術範囲を逸脱するものではない。本発明の顔検出の面がコンピュータプログラムとして実施される場合、プログラムは従来のコンピュータ読み取り可能な記憶媒体に記憶され、これは、例えば磁気ディスク

(例えばフロッピー(登録商標)ディスク)又は磁気テープといった磁気記憶媒体、光ディスク、光テープ又は機械読み取り可能なバーコードといった光学記憶媒体、ランダムアクセスメモリ(RAM)又は読み出し専用メモリ(ROM)といったソリッドステート電子記憶装置、又はコンピュータプログラムを記憶するために使用される任意の他の物理的な装置又は媒体でありうる。

【0014】ここで図1及び図2のブロック図を参照するに、標準カメラの構成要素として実施される一体型システムとしてカメラ10が示され、画像捕捉部20と、プロセッサ又は中央処理装置(CPU)30と、捕捉された画像及び画像に関連する注釈を記憶するデジタルメモリ32と、捕捉された画像及び／又はカメラの動作に有用な他のデータを表示する表示装置34とを含む。捕捉部20は、従来の電荷結合装置(CCD)といったイメージセンサ23上に画像22(本願の目的では1つ以上の顔を含む)を合焦するための自動焦点機能を有する光学部21を含む。露出制御機構24は、イメージセンサ23上への画像の露出を調整するためのアパーチャ及びシャッターを含む。電子捕捉装置の代わりに(又はそれに加えて)、捕捉部は従来の写真フィルムといったアナログ記憶装置25を含む。磁性記録層を含む周知のAPSフィルムの場合、記録装置26は磁性層上に捕捉された画像に関する注釈データを記録しうる。フラッシュユニット27もまた、周囲光が十分でないときに画像22を照明するために設けられる。

【0015】CPU30は、システムバス40を介して、ランダムアクセスメモリ(RAM)42と、読み出し専用メモリ(ROM)44と、(捕捉部20、デジタルメモリ32、記録ユニット26及びフラッシュ27をバス40へ接続するための)入出力(I/O)アダプタ46と、(情報取り扱いシステム又はインターネットといったデータ処理ネットワークへ直接接続するための)通信アダプタ48と、(顔を追跡する測定フレーム49aを生じさせるための)目標追跡段49と、(シャッターボタン52、フラッシュ制御54、プログラムされた露出選択56、ユーザによって操作されるディスプレイカーソル58といったユーザインタフェース装置及び／又は他のユーザインタフェース装置をバス40へ接続するための)ユーザインタフェースアダプタ50と、(顔検出アルゴリズム90を含む様々な記憶されたアルゴリズムをバス40へ接続するための)アルゴリズムインタフェースアダプタ60と、(バス40を表示装置34へ接続するための)表示インタフェース70とに接続される。CPU30は、十分にパワフルであり、顔検出

アルゴリズム90を実行するのに十分な付加メモリ42及び44を有する。

【0016】バス40に接続される学習データベース72は、顔検出アルゴリズム90が非常に広い範囲の撮像条件で動作することを可能にするための十分な学習データを含む。望ましい実施例では、以下詳述するように、顔検出アルゴリズムは2つのコンポーネントアルゴリズムを含む。第1のコンポーネントアルゴリズムは、画像の顔候補領域を顔形状モデルと領域中の画素に割り当てられた顔確率との比較に基づいて推定し、第2のコンポーネントアルゴリズムは画像の各領域を調べるためにパターン解析を用いて顔候補領域に対して動作し、領域中の顔の存在を確認する。この組合せによれば、第1のコンポーネントアルゴリズムは偽の肯定結果が潜在的に存在する可能性があるものの迅速に動作するよう設計され、第2のコンポーネントアルゴリズムはより計算集約的な処理を第1のアルゴリズムを通った比較的少ない領域に制限しうるという利点がある。

【0017】顔検出の結果は、インタフェースアダプタ60を通じてデータバス40に接続されるアルゴリズムに具現化されるカメラの多数の機能を制御するために使用される。顔検出結果は、例えば1つ以上の顔位置の図心を追跡するために測定フレーム49aを設定し操作する目標追跡段49によって追跡される。測定フレームは、ここに参照として組み入れられる米国特許第5,430,809号に記載されるように使用され、顔領域に対する自動焦点、自動露出、自動色調整、及び自動白バランスのために収集されるデータを制限する。測定フレームは、小さいスポット状の領域であってもよく、又は1つ以上の顔の境界と略一致する境界を有するよう構成されてもよいが、いずれの場合もアルゴリズムのために収集されるデータを顔データ又はそのサンプルに制限するためのものである。これらのアルゴリズムは、赤目補正アルゴリズム80、露出制御アルゴリズム82、フラッシュ制御アルゴリズム84、焦点制御アルゴリズム86、カラーバランスアルゴリズム88、及び構図アルゴリズム92を含む。

【0018】赤目補正アルゴリズム80は、フラッシュユニット27によって生ずる赤目条件を除去するための記憶されたデジタル画素値を調整する。露出制御アルゴリズム82は、顔が正しく露出されるよう、画像捕捉部20の画像露出制御機構24のために測定フレーム49aからの設定値を決定する。露出制御の決定と共に、フラッシュアルゴリズム84は、顔画像の最適な捕捉のためにフラッシュを点灯させるべきか否かを判定する。カメラは、最終捕捉画像が顔領域に正しく合焦されるよう、測定フレーム49aからの距離データを導出するため、及びフレーミング画像の結果を用いて光学部21のポイントすることが可能な焦点機構を設定するために焦点アルゴリズム86を使用する。カラーバランスアルゴ

リズム88は、測定フレーム49a中の肌領域の表現を最適化し、それらが肌の色調の予測色範囲と一致するよう、デジタル画像ファイルに対して適用される。

【0019】画像表示装置34は、撮影者が、捕捉前に画像を予め見る（プレビューする）こと及び／又は最後に捕捉された画像を見ることを可能とする。更に、画像をプレビューするために光学ビューファインダ28が設けられる。更に、CPU30は、必要であれば見られた画像中の顔をハイライトするために顔検出アルゴリズムを使用しうる。このために、光学ビューファインダ28中に半透明の液晶ディスプレイ（LCD）オーバーレイ29が設けられてもよい。LCDドライバ29aは、CPU30からの顔位置データに応じてLCDオーバーレイ29の1つ以上の顔位置に対応する幾つかの領域を作動させる（かかるLCDマスクは、ここに参照として組み入れられる米国特許第5,103,254号に開示される）。また、CPU30は、表示装置34上に示されるビュー領域中の顔の上に例えばボックスを表示するために、表示インタフェース70を介してパターン発生器74を駆動することによってハイライトされた又は輪郭をはっきりさせた顔を発生させる。更に、顔は、撮影者が表示装置34のビュー領域上でカーソル58を動かし、これが例えば顔の上に重なるか又は顔の回りにボックスを描くことによってマーク付けされうる。これは、LCDドライバ29aと光学ビューファインダ28中のLCDオーバーレイ29を通じてなされうる。

【0020】本発明の他の利点は、画像中の顔の検出に関連するデータが自動的に記録され、画像の注釈と共に、又は注釈として含まれることである。これは、画像捕捉時又は後の時点に撮影者によって注釈が付けられることを必要とせずに、イベントの写真記録に重要な被写体の自動的に記録することを可能とする。すると、シーン中の顔の検出により、画像捕捉イベントに対して、また続く画像処理に対して、かなりの追加的な改善を行なうことが可能となる。例えば、顔検出は例えば一人以上の人物を被写体として含む画像を取り出すことによる後の検出のために画像をインデックス付けする便利な手段を提供する。従って、顔検出アルゴリズムを実行することにより、1つ以上の検出された顔の位置、向き、スケール及びポーズといった1つ以上のパラメータに対応する顔データが与えられる。

【0021】更に、一旦顔が検出されると、ユーザによってカメラが以前に捕捉し学習データベースに格納された小さい学習顔ギャラリーから顔を識別するために単純な顔認識アルゴリズムが適用されうる。顔の検出及び位置決定の結果は、画像に付された補助データ位置に格納され、これらは共にRAM42内のバッファメモリに格納される。画像には、例えば検出された顔の座標、顔の寸法の推定値、目の位置、頭のポーズパラメータの大きな推定値、及び各個人の識別情報といった注釈が付き

れる。一旦画像が格納のために選択されると、画像及びその注釈データは、デジタル画像メモリ32中に一緒に記憶されるか、又は注釈がアナログ画像メモリ25の磁性層に記憶される（画像は写真乳剤上に従来の潜像として記憶される）。デジタルカメラについての1つの形態では、捕捉された画像データは、記憶媒体中に、シーン中に特定の数の顔を有する画像に専用のデジタルフォルダ中に記録される。

【0022】本発明の更なる実施例は、シーンの画像を、画像センサ23といった電子媒体とAPSフィルム25といったフィルム媒体の両方に同時に捕捉するハイブリッドカメラである。この実施例では、CPU30は、シーン中の1つ以上の顔の存在を決定するために画像センサ23からの画像データを処理し、顔データは、画像中の少なくとも1つの顔の位置、スケール、又はポーズに対応して発生される。かかる顔データは、捕捉された画像を評価するために、ディスプレイ34にカメラのユーザに対して表示されうる。顔データ（又は画像）が捕捉された画像についての問題を示す場合、ユーザはフィルム25の他のフレーム上に画像を再捕捉する機会を有する。更に、記録ユニット26を作動させることにより、フィルム媒体25の磁性層上に顔データを書くことができる。

【0023】図3及び図4に示されるように、カメラはまずフレーミングモードで動作し、次に最終撮像モードで動作する。各モードで、カメラは撮影者を支援するための多数の自動化特徴を提供する。撮影者はユーザインタフェースアダプタ50を通じてフレーミングモードを無効とするオプションを有し、それによりフレーミング画像の捕捉を無効とし、最終撮像モードへ直接進む。

【0024】[フレーミングモード] 図3に示されるフレーミングモードでは、カメラ10は捕捉部20を作動させることにより、ステップ100においてフレーミング画像を獲得する。次に、CPU30はステップ120において顔検出アルゴリズム90を実行し、これはフレーミング画像中に全ての顔を検出しようとし、ビューファインダ28中で又は表示装置34上でそれらの位置を撮影者に対して示す。更に特定のには、顔検出アルゴリズムは、顔を見つけるために学習データベース72からの顔学習データを使用する。決定ブロック130において顔が検出されれば、1つ以上のカメラアルゴリズムの続く使用のために、ステップ140において顔位置データがRAM42中に記憶される。更に、顔位置は表示インタフェース70によって処理され、例えば、表示装置34上に生成される顔はボックス又は他の輪郭付け特徴によって輪郭付けされる。顔検出アルゴリズム90が顔を見つけることが出来なければ、これは決定ブロック130の出力に影響を与える。

【0025】従って、顔検出の失敗に応じて、即ち顔が見つからない場合、撮影者は、参照番号132で示され

る経路を通じて始まりまで戻り、シーンを僅かな間にくり再びポーズさせるか、又は、ビューファインダ28又はディスプレイ34上の顔の位置を手動で決めるためにカーソル58を用いてカメラに対する手動検出入力を与える。例えばタッチ感応スクリーン及びスタイラス（図示せず）といった他の入力技術が使用されうる。次に、顔の存在及びフレーミング画像中についての知識が与えられると、カメラ10は最終捕捉画像を改善させるために使用されうる撮影者にして価値のあるサービスを提供することが可能である。かかるサービスは、以下説明するようなフォーカス支援、露出及びフラッシュ決定及び構図援助手段を含む。

【0026】[フォーカス支援] 最近の多くのカメラは、ビューファインダ中の合焦目標点を用いて自動フォーカシング及びユーザ指定フォーカシングを提供する。人物は、現れるのであれば、通常は画像中の最も重要な被写体であるため、撮影者によって指定されない限りは人物の顔に対して画像を正しく合焦することが理にかなっている。現在、多数の焦点検出領域と、入力刺激（例えば音声命令）に応じて1つの領域を選択するステア可能な選択機構とを含むシステムが公知である（例えばここに参照として組み入れられる米国特許第5,749,000号を参照のこと）。或いは、上述の'809号特許に示されるように、自動焦点は、顔を含むよう設定された測定フレーム中で実行されうる。本発明に関連して、フレーミング画像上で顔検出を実行した後、カメラ10はフォーカス段階150において焦点制御アルゴリズム86に入り、シーン中の多数のシーンについて画像を最適に合焦する特定の焦点検出領域を選択するために光学部21においてそのステア可能な自動焦点システムを使用する。（或いは、焦点は、主要な被写体を構成すると想定されるシーン中の最も大きい顔に対して最適に設定されうる）。

【0027】[露出及びフラッシュ決定] カメラ10は、その露出制御アルゴリズム82及びフラッシュ制御アルゴリズム84を通じて自動露出制御及びフラッシュ使用機能を提供する。典型的なマイクロプロセッサ制御された露出制御装置は、ここに参照として組み入れられる米国特許第4,503,508号に開示され、周囲及びフラッシュ露出の両方に使用される。上記米国特許によって提供される露出制御機能は、上述の'809号特許に関連して説明された測定窓49aの中に配置される顔領域に制限されるか顔領域に偏らされる。人間は通常は人間が現れる画像中で最も重要な被写体であるため、撮影者によって指定されない限りは、人間の顔の見かけが最適となるよう露出を選択することが理にかなっている。フレーミング画像上で顔検出を実行した後、ステップ160においてカメラは、シーン中の多数の顔に対応する検出領域に対して、画像露出を最適に設定するためにその自動露出アルゴリズム82を使用する。（或い

は、焦点は、主要な被写体を構成すると想定されるシーン中の最も大きい顔に対して最適に設定されうる）。同様に、露出制御アルゴリズム82は、シーン中の顔の照明が適切であるかどうかの評価に基づいてフラッシュ27を点灯させるかどうか判定する。照明が適切でなければ、ステップ170においてフラッシュ制御アルゴリズム84はフラッシュユニット27を作動させ制御する。

【0028】[構図援助手段] 顔検出カメラ10は、ステップ180において、撮影者に対して構図アドバイスを与える構図支援モードを提供する。多くの消費者写真は、美しさの観点からは画像構図が良くないという欠点がある。正しくない焦点及び露出と共に、劣った構図は消費者画像印画に対する不満を生じさせる。多数の発見的な「経験則」が、満足のいく画像を生じさせる構図のよい原理として広く受け入れられるようになっている。例えば、小さい主被写体は、しばしば興味深くない印画を与える。また、「経験則」は主被写体が、垂直方向或いは水平方向、又はその両方向に大体画像の3分の1の点に配置されることを必要とする。かかる原理は、Grill, T. and Scanlon, M., *Photographic Composition*, Amphoto Books, 1990に詳細に説明されている。

【0029】顔検出カメラ10は、フレーミング画像中の顔検出の結果に基づいて、構図アルゴリズム92がビューファインダ又は表示装置34に現れる構図案を発生させることを可能とする構図支援モード180を提供する。構図アルゴリズムは、図5に示される段階に従うが、他の構図要素（例えば上述のGrill及びScanlonの文献に記載される）について調べうることも理解されるべきである。1つの援助手段では、検出された最も大きい顔の領域を全体の画像領域と比較する（ステップ181）。比較結果が、顔が小さすぎることを示す閾値を超過する場合（ステップ182）、ディスプレイ34はカメラを主被写体に近づけるよう提案する（ステップ183）。第2の援助手段では、顔の図心と経験則を表わす図6に示される格子線、即ち、主な被写体が最も満足のいくようになる位置、とを比較する（ステップ184）（これについてはGrill and Scanlonの参考文献の第22頁により詳しく記載されている）。顔が実質的に格子線から外れている場合（ステップ185）、ディスプレイ34はより満足のいく画像を達成するために主被写体を経験則に従って配置することを提案する（ステップ186）。

【0030】第3の援助手段は、顔を画像の境界を横切るよう配置する（ステップ187）。閾値が、顔のかなりの量がカメラオーバーチャによって切り取られていることを示す場合（ステップ188）、ディスプレイ34は撮影者に対して警告するよう設定される（ステップ189）。第4の援助手段は、顔の図心を水平な線に関連付ける（ステップ190）。顔が共通の水平な線上にあるようである場合（ステップ192）、ディスプレイ34

はより興味深い画像を生成するよう、人の集まりの画像中の顔の高さを水平に整列させるよりも変化させることを提案する。図5は、かかるアルゴリズム及びここに参照として組み入れられる上述のGrill and Scanlonの参考文献が同様に実施されうることを一般的に示すものである。

【0031】[最終画像モード] 図3に示されるようにフレーミング画像を捕捉し処理した直後、撮影者に対して上述のステップ150乃至180が与えられていれば、カメラは図4に示されるように最終画像を捕捉する準備ができています。図4に示される初めのステップ200乃至240は、図3に示されるステップ100乃至140と同一であるため、更には詳述しない。更に、図4中、更なる援助手段が最終画像に対して直接的に作用する。上述のように、撮影者は、第2の支援手段の群のみが所望であれば最終画像でのみ（フレーミング画像を排除して）動作することを選択しうる。或いは、表示装置34上で見られるフレーミング画像がユーザにとって満足のいくものであると見なされれば、これは永久画像として保存されうる。いずれの場合も、最終画像モードの一部として、以下説明する最適カラーバランス、赤目通知及び補正、向きマーク付け、及び顔ラベル付けを含む幾つかのサービスが提供される。

【0032】[最適カラーバランス] 人間の視覚系は様々なシーン条件に亘って知覚的な恒久性を維持するすばらしい能力を示すが、アナログカメラもデジタルカメラもそれと同じ能力を有するものではない。例えば、人間の目には、対象の色は、対象が太陽光、天空光、又はタングステン光のどれでも同じに見えるが、これらの3つのシーン条件は、単一の種類の写真フィルム上で捕捉された場合は、必ず非常に異なる色を再生することになる。従って、慣習的に、画像を印刷する前又は他の方法で表示する前に、捕捉された画像に対してカラーバランスアルゴリズム(CBA)を適用することが行われている。自動化されたカラーバランスアルゴリズムの現在の技術は、主に最も顕著なシーン光源について補償しようとする。

【0033】画像のカラーバランスは、シーン中の対象の性質についての幾らかの理解とそれらの相対的な重要性を考慮に入れることによってより良く行われる。シーン全体が正しく再生されるが、顔の色がかなり誤っている場合、再生された画像は許容可能でない。再生についての認められる不快さの特質は文化によって異なるものの、人間は肌の色調が正しく再生されていないことについては非常に敏感である。異なる人物及び人種による肌の色調の多様性は、統計的に分類され理解されうる。更に、偶然にも、肌の色の自然の多様性及び肌の色の再生についての不快な誤りは、正しく選択された色空間中の直交方向に存在する傾向がある。肌の色は、青-赤ディメンション上で変化し、許容可能でない再生誤りは主に

緑-マゼンタディメンションに関連する。

【0034】シーン中の顔の存在及び位置についての知識は、2つの異なる方法によって改善されたカラーバランスを与える。(アナログ画像の光学印刷の場合のように) 大局的な画像補正が可能であれば、大局的な照明の推定値は顔の肌の色調を満足いくよう演色するよう調整されうる。顔検出カメラは、アナログフィルム媒体25の磁性層上の顔の位置及び寸法を記録することにより、顔の肌の色調の正しい再生のために後の光学写真仕上げ装置がカラーバランスを最適化することを可能とする。他方、デジタル処理ステップが利用可能であれば、顔領域は照明のより多くの大局的な考察事項とは独立に補正されうる。これは、主要な被写体(人)及び背景領域の両方が満足のいくように再生されうるため、光学手段のみによって得られる最も良い印画を与える最善の可能なシナリオである。いずれの場合も、カメラ10は、少なくとも位置を決められた顔に基づいて画像に対する最善のカラーバランスを与えるために、そのカラーバランスアルゴリズム88を顔優先補正ステップ260において使用する。更に特定的には、CPU30は、検出された顔から色データを補正するため、及び次に顔領域についてカラーバランスアルゴリズムを重み付けするため、追跡段49によって発生される測定フレーム49aと相互作用する。

【0035】[赤目通知及び補正] ここに参照として組み入れられる共通に譲渡される米国特許第5,432,863号に開示されるアルゴリズムのような赤目検出アルゴリズム80は、赤目補正ステップ270において最終捕捉画像に対して実行される。検出された顔の存在は、赤目検出アルゴリズム80において偽の肯定結果の誤りを防止することを支援するための追加的な証拠として使用される。一対の検出された赤目は、顔の存在の証拠を再び強めることによって補強されるべきである。赤目の存在はまた、赤目検出アルゴリズムによって、表示装置34上で適当な警告を指定しうる表示インタフェース70へ与えられうる。赤目通知を受けた後、撮影者は他の画像を得ることを選択しうる。或いは、カメラ10がデジタルカメラであれば、目の中の不快な赤いハイライトを除去するために、自動赤目補正アルゴリズムが呼び出されうる。

【0036】[向きマーク付け] 多くの消費者写真仕上げ注文は、各画像の小さい変形の連続であるインデックスプリントと共に戻される。画像の全てが正しい自然な向きで印刷されていなければ、インデックスプリントの有用性は低下する。画像中の顔の存在は、その正しい向きに関して強力な手がかりを与える。例えば、顔のディメンションは、2つの主な軸へ分割され、最も長い軸が上下軸として扱われることができる。次に以下説明する顔検出アルゴリズムのうちの1つは毛髪領域を区別し、それにより直立の向きを示す。顔の大多数は、全体的な

画像の向きからみて、直立であるか直立に近い。顔検出カメラ10の顔検出アルゴリズム90は、向きを検出し、画像記憶装置32（又は25）上の各捕捉された画像に対して向き決定ステップ280において検出された顔の向きによって提案される正しい向きの注釈を付す。

【0037】「顔ラベル付け」一旦顔が検出されると、ユーザの操作によってカメラが以前に捕捉しており学習データベース72に記憶されている小さな学習顔ギャラリーから顔を識別するために、単純な顔認識アルゴリズムが適用される。ギャラリーは、例えば家族の一人一人、又は学級の子供たちを含む。カメラによって新しい画像が捕捉されており、顔が検出されると、顔認識アルゴリズムによって確立される各個人の識別情報は、画像と共に、デジタル記憶装置32又はフィルム25の磁性層に記録される。かかる顔識別情報は、画像と共に、写真仕上げ段階又は続くコンピュータ処理段階へ渡される。情報は、画像中の人の名前を印画に自動的にラベル付けするために使用される。他の可能な適用は、単一の人物又は顧客によって指定される人物の集まりの画像を含むアルバムを自動的に生成することである。典型的な顔認識アルゴリズムについて、それらのアルゴリズムがより大きいシステムとして埋め込まれることを可能とするソフトウェア開発キットを提供する多数の市販の顔認識製品がある。例えば、Visionics社製の「Face-It」システムは、顔認識アルゴリズムとしての使用に適している。

【0038】「顔検出アルゴリズム」デジタルカメラ中で動作する顔検出アルゴリズムは、所与の限られた計算計算及びコンピュータメモリ資源で成功するのに必要な規準を満たさねばならない。即ち、アルゴリズムは高速に（即ち1秒以下）で、真の肯定結果／偽の肯定結果の検出率について十分に高い性能を有さねばならない。限られた資源について埋め合わせるように、アルゴリズムの結果が人間の操作者へ示されること又は人間の操作者によって使用されることはアルゴリズムの失敗に対してある程度の耐性があることを意味する。この耐性は、本願の発明の特徴を可能とするものである。

【0039】本実施例では、組み合わせられて使用されることにより検出レベルに関しては個々のアルゴリズムよりも高い性能を与える2つの顔検出アルゴリズムの組合せを使用することを提案する。第1の検出オペレータ、即ちコンポーネントWは、顔候補についての非常に高速な前スクリーン部である。第2の検出オペレータ、即ちコンポーネントSは、偽の肯定結果の非常に遅い速度によって特徴付けられる洗練されたパターンマッチングアルゴリズムである。コンポーネントWによってラベル付けされた顔候補は、続いてコンポーネントSによって検査され、最終検出決定が与えられる。

【0040】「コンポーネントW」Wu外は、デジタルカメラに組み込まれるのに非常に良く適した顔検出アル

ゴリズム（以下、変更されたように、コンポーネントWと称する）について発表した（ここに参照として組み入れられるWu, H., Chen, Q., and Yachida, M., "Face Detection from Color Images Using a Fuzzy Pattern Matching Method", IEEE Trans. Pattern Analysis and Machine Intelligence, 21 (6), 557-563, 1999参照）。アルゴリズムは非常に高速であり、プログラムメモリ及び学習状態の両方について非常に少ない量のみを必要とする。コンポーネントWは、色特徴に基づいて顔を含む可能性の高い画像窓について探索するある種のアドホックなパターン認識オペレータである。方法は本質的には、色及び形状に基づいて肌を含む可能性が高いと思われる中央部と、やはり色及び形状に基づいて毛髪を含む可能性が高いと思われる周辺領域（肌の上方及び側方の周り）を有する窓を探す。方法は色信号に基づくため、方法が動作する画像は有意味な測色法で符号化されることが必要である。

【0041】コンポーネントWは、学習フェーズと試験フェーズを有する。学習フェーズは、肌及び毛髪の色分布の収集と、適当な学習例から形状モデルを集めることからなる。試験フェーズでは、窓は、画像全体に、スケール及び位置の全ての範囲で走査される。コンポーネントWは、画像が上向きであることが知られており顔が大まかに画像の向きと整列されていることを暗に仮定している。この仮定は、各可能な画像の向きについて一回、顔全体の探索を数回（カメラは上下には使用されないため望ましくは3回）実行することによって緩められる。試験フェーズでは、アルゴリズムは検査されるべき各画像窓について一回、以下のステップを適用する。

【0042】（1）肌及び毛髪の高率マップを計算する。

【0043】デジタル画像中の各画素を肌及び毛髪の色所定の確率テーブルと比較し、画素が人間の肌又は毛髪を表わす事後確率を与える。確率テーブルは、オフラインで収集され、カメラ内に格納されねばならない。これらは、同一の分光感度を用いて、デジタルカメラと同じ撮像センサによって収集される。

【0044】（2）非線形オペレータを介して確率を推定領域分数へ変換する。

【0045】顔形状モデルは、やはりオフラインで学習例から構築される。これらのモデルは、標準的な頭のポーズの小さい組の中の空間的に正規化された人間の顔に重ね合わされた矩形のグリッドの各セル中に肌及び毛髪の色が発生する尤度を符号化する。

【0046】（3）顔形状モデルを用いてファジーパターンマッチングを実行

画像の各画素位置へ矩形の窓が順番に走査され、窓が顔を含むかどうかについて判定が行われる。様々な寸法の顔に適用するために、走査プロセスはある範囲の寸法に亘って異なる窓を用いて繰り返される。顔が画像の窓の

中に存在するかどうかの判定は、所定の顔形状モデルと窓の各セル中の肌及び毛髪の後確率の実際の分布とを比較することに基づく。ファジー比較は、Wu外の文献に記載されるように、最善の結果を与えるために較正段において調整されたパラメータ化された非線形オペレータを使用する。

【0047】以下、顔形状モデルについて説明した後、上述の各ステップについて詳述する。更なる詳細についてはWu外の文献を参照することによって分かることが理解されるべきである。

【0048】[形状モデル] 顔形状モデルは、典型的な顔のポーズにおける肌及び毛髪の空間的な分布の低解像度表現である。各別個のポーズには肌用の1つのモデルと毛髪用の1つのモデルがある。各モデルは $m \times n$ 個のセルからなり（ここでは、 $m=12$ であり、 $n=10$ ）、各セルは所与のポーズにある典型的な顔部について、（肌モデルについては）肌によって占められるセルの割合、又は（毛髪モデルについては）毛髪によって占められるセルの割合を符号化する。画像窓は、形状モデル中のセルに対して空間的に対応付けられうる。窓の解像度に依存して、単一の画素又は画素のブロックは、各モデルセルに対応しうる。モデルは、2つの目を標準的な位置に配置するため、アフィン変換が適用された一組の学習画像を用いて構築される。空間的に正規化された画像は、次に肌領域及び毛髪領域へ手動でセグメント化され、モデルの解像度における割合的なセルの占有度が計算される。図7A乃至7Dは、正面ポーズ及び準正面ポーズについての形状モデルの例を示す図である。モデルは、占有度割合のグレーレベル符号化と共に学習データベース72（図1参照）に格納される。

【0049】[毛髪及び肌確率を計算] この時点では、学習画像から肌の色及び毛髪の色についての確率分布を獲得することが目的である。目標は、肌をskin、毛髪をhair、色をcolorとしたとき、 $P(\text{skin} | \text{color})$ 及び $P(\text{hair} | \text{color})$ の形状の確率テーブルを獲得することである。Wu外の文献ではFarnsworthの知覚的に均一な色空間が用いられるのに対して、本発明では、肌領域及び毛髪領域を区別するために望ましい測色法として（L, s, t）色空間を使用し、従って、（L, s, t）測色法で確率学習及び適用を実行し、但し、 $L = c(R + G + B)$; $s = a(R - b)$; $t = b(R - 2G + B)$; a, b, c は定数; R, G, B は相対対数シーン露出に比例する画像値である。この測色法は、画像セグメント化を実行すべき有効な色空間であることが証明された。3つの全てのチャンネルが使用されるが、確率ヒストグラム中でルミナンスチャンネルは組み合わされたクロミナンスチャンネルから分離される。

【0050】肌の色の統計値を集めるため、1800程度の画像を含む注釈の付けられたデータベースが使用され、各画像は12ビットの相対対数シーン露出RGB測

色法で記憶される。これらの画像のためには、人間によって判定されるカラーバランス及び白い点の目標と、2つの目が見える場合はデータベース中の全ての顔の目の位置とが利用可能である。人体測定学的に平均的な顔モデルを用いて、データベース中の全ての顔について顔の肌が抽出される。カラーバランス及び白い点の目標もまた、夫々の場合において画像から差し引かれる。画素値は次に以下のマトリックス計算、

【0051】

【数1】

$$\begin{pmatrix} L \\ s \\ t \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{6}} \end{pmatrix} \begin{pmatrix} \bar{r} \\ \bar{g} \\ \bar{b} \end{pmatrix}$$

を用いて（L, s, t）測色法へ変換され、式中、ハット付きの量は目標値について調整されている。毛髪の色は統計値を集めるために、各例として顔部について毛髪領域が手動でセグメント化されることを除き、従来同様のプロセスが実行される。肌又は毛髪の各画素は、

（L, s, t）空間中の色の例を与える。毛髪及び肌の両方について別々の確率ヒストグラムが累積される。Lヒストグラムは、夫々の場合に2次元のs, tヒストグラムから別々に作成される。このように、肌及び毛髪の色及びルミナンスが別々に分布するという暗黙の仮定が取られる。試験時に、画像はまず各画素における色値を用いてその画素に対して肌又は毛髪のいずれかがイメージングされている事後的な尤度を参照するために処理される。ワールド対象中の色の分布が考慮に入れられることを確実とするため、確率分布に対してベイズの定理が適用される。この計算の結果、図8A及び8Bと、図9A及び9Bに示されるような肌及び毛髪確率密度の2つのグラフィックな表示が与えられ、各グラフィックな表示は、ルミナンス成分とクロミナンス成分の間で分離された肌画素確率及び毛髪画素確率の夫々の2つの図を表わす。

【0052】[肌確率及び毛髪確率を推定された面積割合に対して写像する] 形状モデルは、顔画像中の色の空間的な分布に関する情報を含み、一方、以前の段階において計算された確率は候補顔領域中の平均の色に依存する。これらは異なるため、2つの情報カテゴリを直接比較することはできない。従って、非線形写像関数を使用して画像窓中の確率値を推定された肌占有割合及び毛髪占有割合へ写像するために中間的な段階が必要とされる。非線形関数は、調整パラメータa及びbを有するS字状関数であり、以下の式、

【0053】

【数2】

$$S(x) = \begin{cases} 0 & x \leq a \\ \frac{2(x-a)^2}{(b-a)^2} & a < x \leq \frac{(a+b)}{2} \\ \frac{2(x-a)^2}{(b-a)^2} & \frac{(a+b)}{2} < x \leq b \\ 1 & b < x \end{cases}$$

によって与えられる。

【0054】Wu外の文献は、最善の顔検出結果を生じさせるために経験的な実験に基づいて肌モデル及び毛髪モデルの夫々に対してパラメータa及びbを別々に調整することを記載している。本発明では、写像の非線形オペレータは、肌色及び毛髪色の統計的なサンプリングの欠陥を補償するのに最も有用であることが分かった。このために、パラメータa及びbはかなり低く設定される。目標は、コンポーネントWが殆ど全く顔を見逃さず、コンポーネントWを通る多くの偽の検出を除去するためにアルゴリズムaによって更なる処理を行なうことである。

【0055】「顔形状モデルを用いてファジーパターンマッチングを実行する」各セルについての肌及び面積占有割合を有する形状モデルと、非線形写像から生ずる対応する画像領域の同じ量の推定値とが与えられているとき、画像領域とモデルセルの間の類似性についての判定が行なわれる。類似性の尺度は、二項「ファジー関係式」、即ち、

【0056】

【数3】

$$\text{similarity}(I_s, I_h, M_s, M_h) = e^{-a \cdot \text{dist}(I_s, I_h, M_s, M_h)^b}$$

$$\text{dist}(I_s, I_h, M_s, M_h) = \sqrt{(I_s - M_s)^2 + (I_h - M_h)^2}$$

を使用し、式中、パラメータa及びb（上記のセクションの非線形写像のパラメータとは異なる）は、比較の形状を決定し、I及びMに添え字が付された変数は夫々、画像領域及びモデルセルからの肌確率及び形状確率を夫々表わす。bを増加させることにより、形状モデルと画像窓のとの間の差に指数的により急なペナルティを与え

P (face | image)

に対してなされるここに記載される単純化は、大きな無数の分布を非常に実際的なものへ変化させる。目的は、単純化された分布P (face | distilled-image-feature)を得て、それにより抽出された画像特徴 (distilled-image-feature) が計数され、学習中に例えば100万binへまとめられることである。学習分類オペレータの発見的な方法は、binよりも2つのオーダだけ大きい例が必要とされることを示す。従って、10⁶のbinを、統計学的に有意な方法で、ある場所を占めさせるためには、10⁸の例が必要でありうる。特にそれらのうちの幾つかを計算的に発生させることが可能である場合、及び例の「単位」が人間の顔全体よりも小さいも

る。本実施例では、幾らかの実験後、値a=2及びb=1が選択された。画像窓全体についての類似性のスコアは、形状モデルの全てのセルに亘る平均類似性スコアとして扱われる。コンポーネントWによって検出される顔候補を識別するために同様の尺度に対して閾値が適用されうる。

【0057】「コンポーネントS」Schneidermanアルゴリズム（以下、コンポーネントSと称する）についての詳細な説明は、ここに参照として組み入れられるSchneiderman, H. and Kanade., "Probabilistic Modeling of Local Appearance and Spatial Relationships for Object Recognition", Proc. CVPR 1998, 45-51に記載されている。方法の主な段階について、独立した説明を与えるため及び参考文献に対する差異及び改善点について強調するために、ここで概説する。コンポーネントSは、faceは顔、imageは画像であるとする、条件付き確率分布P (face | image)を近似する記憶された確率分布を用いて最大の事後的な分類を行なうベイズ分類オペレータを実施する。方法は、画像からの証拠が存在する場合に事前に測定された学習分布P (image | face)を事後的な分布へ変換するためにベイズの原理を使用するため、ベイズ法と称される。証拠は、空間的及び強度的に正規化された画像窓中の画素値からなる。ベイズの定理の使用は、画像証拠が曖昧であってもよいという観察により義務付けられる。幾つかの場合は、関心となるクラス（例えばこのコンテキストでは顔）にない対象及びシーンは、クラス（＝顔）の対象と間違えられうる画像パターンを生じさせうる。ベイズの原理は、「ワールド」画像として知られる代表的な非クラス画像の集合を必要とする。ワールド画像の集合は、アルゴリズムを学習する場合に最も困難且つ重要なプロセスであると分かる。困難性は、ワールドが非常に多様であるという事実に基づく。顔は（少なくともワールドと比較される場合は）多様ではなく、顔の学習例の収集は非常に明白に行われる。この困難性について、以下のセクションにおいて学習について詳細に説明する。

【0058】以下の分布、

(1)

のである場合は、多くの例を収集することが可能である。

【0059】単純化は、以下の順序で適用され、以下のセクションにおいて説明される。

【0060】(1) 顔領域寸法を標準化する。

【0061】(2) 顔領域を副領域へ分解する。

【0062】(3) 副領域間の依存性を無視する。

【0063】(4) PCAを用いて副領域をより低い次元の表現へ投影する。

【0064】(5) 疎係数を用いて投影を符号化する。

【0065】(6) 疎係数を量子化する。

【0066】(7) 外観及び位置を分解する。

【0067】(8) 非共通パターンについての位置を無視する。

【0068】(9) 共通のパターンについて位置依存性をベクトル量子化する。

【0069】(10) 解像度間の独立性を想定して

$$P(\text{face} | \text{region})$$

に変化し、但し、顔はface、領域はregionとすると、領域はちょうど、56×56の画素の画像窓からの画素のラスタ化されたベクトルである。

【0072】2. 顔領域を副領域へ分解する。

【0073】各顔領域は、多数の重なり合う16×16画素の副領域へ分解される。これらの副領域は、領域中の各画素位置、又はこれらの位置のサブセットに固定される。ここでは、副領域を3つ目のライン毎に3つ目の画素に固定する。このように選択することにより、顔領域中には副領域についての196個の可能な固定位置があり、従ってこの位置は単一のバイト中へ符号化される。式(2)の右辺において、「領域」は「副領域」、即ち副領域の集合によって置き換えられうる。副領域の寸法は、適切に配置された場合は、個々の副領域が顔特徴(例えば目、鼻、又は口)を含むのに十分に

$$\prod_{i=1}^{\# \text{subregions}} P(\text{face} | \text{subregion}_i)$$

で置き換えられ、統計的な独立性は多数の副領域についての結合依存性がないことに反映される。

【0077】4. 主要成分解析(PCA)を用いて副領域をより低い次元の表現へ投影する。

【0078】副領域は、夫々に対して256のグレースケールを有する256の画素を含むため、可能な副領域の数は非常に多い。次の簡化は、副領域の次元性を256から12へ減少させるために線形PCAの標準的な技術を適用することを含む。(12の次元の選択はある程度は任意である。実際の副領域の分散の90%以上は12以下の次元を用いて符号化される。)PCAを実行するために、全ての副領域がデータ解析に関係するよう、顔画像の大きな学習組が処理される。異なる画像解像度及び多数の顔ポーズについて別々の主要な成分が必

$$[\text{proj}] = A^T [\text{subregion}]$$

によって実行され、式中、projは投影、subregionは副領域、Aは学習副領域の固有ベクトル(主要成分)を含む列を有する投影行列である。PCAは顔画像の学習組に対してのみ作用することに留意されたい。結果としての主要成分は可能な顔でない画像の非常に大きい組を統

$$\prod_{i=1}^{\# \text{subregions}} P(\text{face} | \text{proj}_i)$$

となる。

【0081】5. 疎係数を用いて投影を符号化する。

【0082】副領域表現は、12個全ての投影係数を保

(1)乃至(9)を多数の解像度に適用する。

【0070】1. 顔領域寸法を標準化する。

【0071】空間的に正規化された顔は、56×56画素の窓の中に表わされる。この簡化により、式(1)は、

(2)

大きくなるよう選択される。この寸法は、整合した単位として調べられうる最も大きい画像特徴を制限する。

【0074】3. 副領域間の依存性を無視する。

【0075】副領域間の統計的な依存性のモデル化は行なわれない。従って、この簡化はモデル化されうる対象の属性のタイプを制限する。例えば、単一の副領域は一つの目を捕捉するのに十分な画素を含みうるのに対して、他方の目は異なる副領域にあり、2つの目についての類似性又は非類似性については考慮されない。いかなる理由付けも顔の異なる部分の照明の相対的なレベルに基づくことは出来ない。この簡化を用いて、式(2)は、

【0076】

【数4】

(3)

要とされるか否かを調べるに、幾つかの実験が行なわれた。これらの結果に基づき、各解像度に対して主要成分の別々の組が記憶されるが、顔のポーズによって異なる組を保持する必要はないことが決定された。直感的に、異なる解像度において本質的な顔構造はユニークな空間パターンを与え、他方で、僅かに異なる顔のポーズによって生ずる変化は最初の幾つかの主要な変化態様ではあまり重要でないことがわかる。

【0079】投影段階により、各画像副領域は主要成分軸に沿った12個の投影係数によって表わされる。この表現は、各副領域を12個の主要な副領域の線形の組合せによって表わすこととなる。投影演算は、以下の行列演算、

(4)

計的に不適当にサンプリングすることによって生ずる幅広い多様性を受ける可能性が高いため、偽(顔でない)例は使用されない。この段階の結果、式(3)は、

【0080】

【数5】

(5)

持するのではなく、6つの上位の係数のみを保持することによって更に圧縮される。しかしながら、この疎符号化スキームは、少なくとも6つの係数を対としてグルー

ブへとグループ化すること、及び保持すべき6つの投影次元を選択する時にそれらの合計の平方値を考慮することによって更に複雑にされる。このようにして、12個の係数は続く処理のために6つへ減少される。

【0083】6. 疎係数を量子化する。

【0084】副領域表現の更なる圧縮は、ロイド・マックス量子化器を用いた9個の係数の別々の量子化を通じて行なわれる。この量子化器は、独立変数のガウス分布の想定の下で平均平方量子化誤差を最小化する。多数の量子化値についての共通値については、bin中止点及びロイド・マックス量子化器の再構成レベルは、テーブル化され、これはLim, J., *Two-Dimensional Signal and Image Processing*, Prentice-Hall: New Jersey, 1990に記載されている。ガウス分布仮定の有効性を試験するため、学習組の投影係数の実際の分布が収集され、そこからガウス仮定が実際の分布に非常によく一致することが分かった。

【0085】保持される疎係数の数及び各係数に割り当てられる量子化レベルの数の選択は、画像副領域を符号化する可能な量子化値の数を決定する。各次元に対して8つ、4つ又は2つの量子化レベルを選択して、6つの顕著な次元を選択することにより、アルゴリズムは実施されると約1,000,000の数のうちの1つによって各副領域を表わしうる。これらの量子化された数は、

$$\prod_{i=1}^{n_{subregions}} P(face | q1_i) \quad (6)$$

へ簡単化される。

【0088】7. 外観及び位置を分解する。

【0089】確率分布の簡単化の流れのこの点において、式(6)は副領域の画素パターンと顔領域内のその

$$\prod_{i=1}^{n_{subregions}} P(face | q1, pos_i) \quad (7)$$

によって置き換えられ、各副領域はここで量子化値と顔領域中での位置(pos)によって表わされる。式(7)の解釈は、直感的に以下の考えを生じさせる。即ち、目のようなパターンは目を含む可能性の高い副領域中のみ生ずるはずである。

【0091】8. 共通でないパターンについての位置を無視する。

【0092】各副領域に対して1,000,000個の量子化レベル及び196個の位置が可能であるとき、式(7)に対する更なる簡単化がなされなければならない。この式に対して2つの更なる簡単化がなされる。第一に、最も共通に生ずるq1パターンのみの位置的依存

$$\prod_{i=1}^{n_{subregions}} \frac{P(face | q1_i)}{n_{pos}} \quad (8)$$

となり、式中、 $n_{pos} = 196$ は可能な副領域の位置

Schneiderman外の文献においていくらか不可解に「q1」値と称されるものである。可能なq1値の数は、その文献において「 n_{q1} 」と称されるアルゴリズムサイズ変更パラメータである。

【0086】この量子化スキームの圧縮の利点は、256²⁵⁶の可能な副領域パターンが10⁶の別々の数へ符号化される場合に明らかとなる。実際、この量子化スキームを画像符号化の形状として考えることが可能である。画像をその符号化物から再構成することは、下の画像のある種の近似を与える。図10は、下の画像と、PCA投影の後、及び疎符号化及び量子化の後の再構成を示す図である。更に特定的には、図10Aは下の画像を示し、図10Bは12の次元の主要成分空間への副領域の投影からの再構成を示し、図10Cは図10Bの疎符号化され量子化された変形からの再構成を示す。(図10B及び図10Cの画像は全ての符号化された情報を示すものではないことに留意されたい。これらはむしろ、副領域が56×56の顔領域の傾斜格子と整列された符号化からの再構成を示す。同時の符号化は、副領域が領域格子に対してずれるにつれ、更なる画像情報を捕捉する。)量子化段階に続き、確率の式(5)は、

【0087】

【数6】

位置の両方を明示的に含むよう拡張される。式(6)は、

【0090】

【数7】

性を符号化するための決定が行なわれる。このため、大きなソート段階は、学習組における発生の頻度を減少させることによりq1パターンを順序付ける。発生閾値以下にソートされる全てのq1パターンの位置的な依存性は均一な位置的に分布によって置き換えられる。学習中に明示的に学習されるべき位置的分布を有するq1パターンの数は、Schneidermanの文献中「 n_{cat} 」と称されるアルゴリズムサイズ変更パラメータである。共通でないパターンについては、式(7)は、

【0093】

【数8】

の数である。

【0094】9. 共通パターンについての位置的依存性をベクトル量子化する。

【0095】式(7)に対する第二の簡単化は、学習中に学習される位置的分布の数を更に減少させることを含む。本明細書のセクション8では、既に、学習される位置的分布の数が n_{q1} から n_{est} へ減少されている。ここで、 n_{q1} から n_{est} への更なる減少は、 n_{est} 生存位置的分布を n_{q2} へベクトル量子化することによって実行される。この量子化のために、 $q1$ パターンの二次元の位置的分布がベクトルへラスタ化される。数 n_{q2} はアルゴリズムサイズ変更パラメータである。

【0096】ベクトル量子化学学習アルゴリズムは、標準的なLBGアルゴリズムではないが、むしろ、入力ベクトルを一回通ることによって実行されるアドホックなカスタムアルゴリズムである。学習アルゴリズムは数万又は数十万のベクトルを量子化することが高く、従って速度に関して考慮に入れねばならないため、この一回通過性質は重要である。学習プロセスは、以下のように概説される。

【0097】各ベクトル x について、最も近い現在のパターンの中心を見つける。

【0098】 x と最も近い中心との間の距離 d を計算す

$$\prod_{i=1}^{\#subregions} P(face | q1, pos_i') \quad (9)$$

となる。

【0104】10. 解像度間の独立性を仮定して多数の解像度において検出を適用する。

【0105】副領域間の統計的依存性は開発された簡単化された確率モデルの中に捕捉できないため、副領域よりも大きい特徴について考慮することができない。この

$$\prod_{j=1}^{n_{mags}} \prod_{i=1}^{n_{subs}} P(face | q1_i^j) \quad (10)$$

を与える。典型的な例は、画素の解像度の $n_{mags} = 3$ レベルにおいて捕捉された単一の顔のものである。各解像度において、目は標準的な位置になくならない。

〔簡単化された確率分布の完全な形成〕式(8)及び(10)を合わせ、これらの式中の事後確率に対して学

$$P(face | region) = \prod_{j=1}^{n_{mags}} \prod_{i=1}^{n_{subs}} P(q1_i^j | face) \frac{P(pos' : q1_i^j, face) P(face)}{\frac{P(q1_i^j | face)}{n_{pos}} P(face) + \frac{P(q1_i^j | \overline{face})}{n_{pos}} P(\overline{face})} \quad (11)$$

式中、

【0108】

【外1】

$P(face)$ 及び $P(\overline{face})$

る。和平方誤差 (SSE) 法が使用される。

【0099】 $d <$ 閾値であれば、 x をクラスタに追加し、クラスタ中心を更新する。

【0100】そうでなければ、 x をシードとして新しいクラスタを始める。

【0101】このアルゴリズムが正しく機能するためには、もちろん空のクラスタを扱い、クラスタの最大数 n_{q2} の位置が正しくないことについても扱わねばならない。クラスタ中心は、クラスタを写像するベクトルの平均として計算される (理想的には、発生計数の加重平均)。距離閾値の選択には問題があり、本質的には異なる値の閾値を使用するときの量子化学学習の挙動を経験的に観察することに基づく。この選択の目標は、利用可能な量子化レベルの数をできるだけ利用すると共に、分布ベクトルを出来るだけ均一に広げることである。

【0102】位置的分布のベクトル量子化を適用するとき、式(7)中の位置 pos は、 pos' と示される VQ パターン中心のうちの1つへ写像される。次に、式(7)はより共通のパターン、

【0103】

【数9】

限界を克服するため、ここで、多数のレベルの画像解像度を導入する。式(2)中の確率推定オペレータの全体の機構は、画像解像度の多数のレベルに対して適用され、

【0106】

【数10】

習中に合わされた事前確率に関するベイズの定理を適用することにより、画像領域中の顔の存在の推定される尤度の完全な形式(11)が生ずる。この式の完全な導出の詳細は、Schneidermanの文献に記載されている。

【0107】

【数11】

は、画像領域が顔を含む事後確率又は含まない事後確率を表わす。これについての知識がない場合、 $1/2$ の均一な事前確率が使用され、上述の式(1)を更に簡単化する。事前確率についてのこの仮定は、顔のパターン認識に使用された場合のアルゴリズムの性能に影響を与え

ない。むしろ、確率値として出力されるアルゴリズムを解釈するときに考慮されねばならない拡大縮小率の存在を与える。

【0109】〔学習段階－フェーズⅠ〕アルゴリズムSの実際の学習は多数の別個の段階を含み、学習は自然に2つの主なフェーズへ分割される。第一のフェーズの目標は、顔副領域の量子化の特定パラメータを獲得することである。最初の段階は、学習組から共分散行列を捕捉し次に副領域の主要成分を捕捉することである。この段階の一部として、主要成分の抽出に続き、12個の主要な次元への投影の統計を集めるため、全ての学習副領域を通る第二の通過が行なわれる。次に投影データが解析される。投影統計は、ロイド／マックス量子化器の最適な設計を可能とするため、学習プログラムへフィードバックされる。顔パターンの変化は、様々なスケールの解像度に亘って考慮され、主要成分及びこれらの成分に沿った学習データの統計的な分布の抽出は、各画像解像度について繰り返されねばならない。

【0110】〔学習段階－フェーズⅡ〕学習の第2のフェーズは、学習組を通過し、各顔例の各副領域の量子化を実行することによって開始する。上述のように、学習組は、各学習典型の僅かに乱れた変形を形成することによって拡張されうる。量子化された値が出現する頻度は、大体1,000,000binを有するヒストグラム中で計数される。同様に、各量子化された値が生ずる副領域位置は累積される。ソート操作により、量子化頻度ヒストグラムは発生計数の降順に配置される。 n_{est} の最も周波数量子化されたパターンでは、位置的な分布はベクトル量子化アルゴリズムに入る。ベクトル量子化に続き、 n_{q2} の種の位置的分布が保持され、各 n_{est} の頻度量子化値は保持される分布によって近似される位置的な分布を有する。

【0111】〔顔検出オペレータを適用する〕学習された顔検出アルゴリズムを試験時に使用することにより、式(11)の計算は空間的及び強度的な正規化が行なわれた画像領域に対して適用されねばならない。各候補顔領域の三つの異なる解像度の変形が必要とされる。各副領域についての量子化値が計算され、式(11)中の様々な確率項はアルゴリズム学習中に形成された確率テーブルから抽出される。

【0112】顔検出のために式(11)を使用するために、確率閾値が選択されねばならない。事後確率が閾値を超過する場合、顔検出が生じている。アルゴリズム学習プロセスが完了した後、顔画像及び顔でない画像の確認組に対して適用された場合のアルゴリズムの分類性能を調べることによって決定される。閾値は、偽の肯定結果の誤りと偽の否定結果の誤りの相対的な重要性を考慮して、確認組に対して最適な性能に設定される。

【0113】〔二重スクリーニング顔検出オペレーター組合せアルゴリズム〕本発明の望ましい顔検出アルゴリ

ズムでは、コンポーネントWによって発生された顔候補はコンポーネントSの顔検出オペレータへの入力となる。2つのアルゴリズムの窓形状は僅かに異なるため、顔候補をフレーミングし、コンポーネントSのために標準的な位置に目を配置するために、空間アフィン変換が使用される。画像窓中の顔の存在を宣言するために、コンポーネントSからの出力に対して閾値が適用される。

【0114】コンポーネントWは画像全体に亘って走査された窓のある範囲のスケールにおいて画像を調べるため、真の顔が一つ以上のスケールで、及び幾つかの近接して離間した窓位置において検出される可能性がある。検出の重なり合いを組み合わせるための幾つかの方法が使用されるべきである。二つの異なる方法が試験された。第一の方法は、空間的に重なり合う検出の群から最も強い検出を使用するものである。第二の方法は、重なり合う検出の平均の目の位置を計算することである。平均化技術は、人間の観察者によって目で判断されるようにより正確な目の位置が得られることが経験的に分かった。

【0115】

【発明の効果】本発明の各面によれば、ユーザに対する写真撮影経験を改善するため、及び撮影者がより良く、より満足のいく写真を得るための多数の提案を提供するために、カメラが顔検出機能を使用するという利点がある。

【図面の簡単な説明】

【図1】本発明によるカメラ要素の配置を示す顔検出カメラについてのブロック図である。

【図2】図1に示されるカメラの画像捕捉部を示すブロック図である。

【図3】図1に示されるカメラがフレーミング画像モードにあるときの動作に含まれるカメラ動作を示すフローチャートである。

【図4】図1に示されるカメラが最終画像モードにあるときの動作に含まれるカメラ動作を示すフローチャートである。

【図5】構図案を発生させる方法を示すフローチャートである。

【図6】経験則を適用するために格子へ分割される画像領域を示す図である。

【図7A】顔検出アルゴリズムのうちの1つで使用される正面ポーズについての形状モデルの例を示す図である。

【図7B】顔検出アルゴリズムのうちの1つで使用される正面ポーズについての形状モデルの例を示す図である。

【図7C】顔検出アルゴリズムのうちの1つで使用される右からの準正面ポーズについての形状モデルの例を示す図である。

【図7D】顔検出アルゴリズムのうちの1つで使用され

る右からの準正面ポーズについての形状モデルの例を示す図である。

【図8A】顔検出アルゴリズムのうちの1つで使用される肌についての確率密度を示すグラフ表現である。

【図8B】顔検出アルゴリズムのうちの1つで使用される肌についての確率密度を示すグラフ表現である。

【図9A】顔検出アルゴリズムのうちの1つで使用される毛髪についての確率密度を示すグラフ表現である。

【図9B】顔検出アルゴリズムのうちの1つで使用される毛髪についての確率密度を示すグラフ表現である。

【図10A】元の画像を示す図である。

【図10B】顔検出アルゴリズムのうちの1つによる主要成分解析に続く再構成を示す図である。

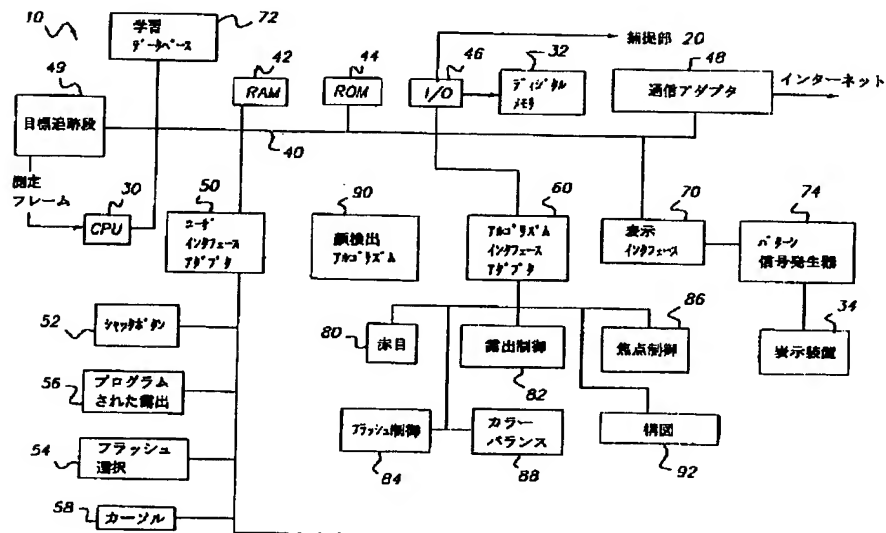
【図10C】顔検出アルゴリズムのうちの1つによる主要成分解析に続く再構成を示す図である。

【符号の説明】

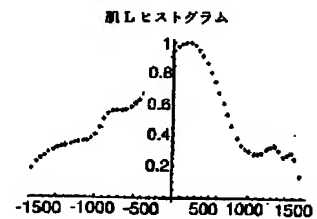
10 カメラ
20 画像捕捉部
21 光学部
22 画像
23 画像センサ
24 露出制御機構
25 アナログ記憶装置
26 記録装置
27 フラッシュ
28 光学ビューファインダ
29 LCDオーバーレイ

29a LCDドライバ
30 CPU
32 デジタルメモリ
34 表示装置
40 システムバス
42 RAM
44 ROM
46 I/Oアダプタ
48 通信アダプタ
49 目標追跡段
49a 測定フレーム
50 ユーザインタフェースアダプタ
52 シャッターボタン
54 フラッシュ制御
56 プログラムされた露出選択
58 ユーザ操作された表示カーソル
60 アルゴリズムインタフェースアダプタ
70 表示インタフェース
72 学習データベース
74 パターン信号発生器
80 赤目補正アルゴリズム
82 露出制御アルゴリズム
84 フラッシュ制御アルゴリズム
86 焦点制御アルゴリズム
88 カラーバランスアルゴリズム
90 顔検出アルゴリズム
92 構図アルゴリズム

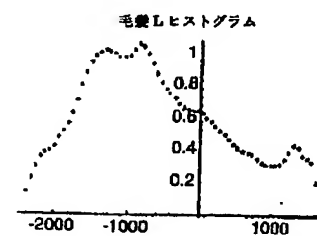
【図1】



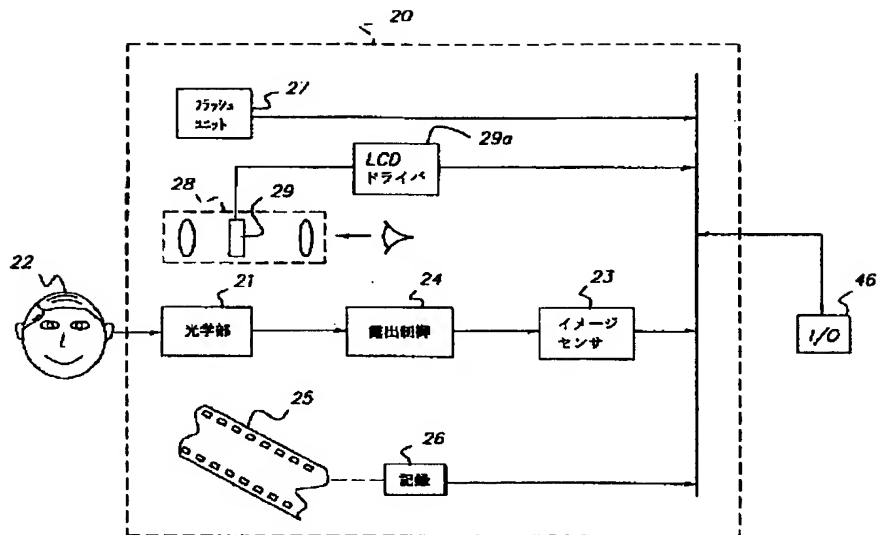
【図8B】



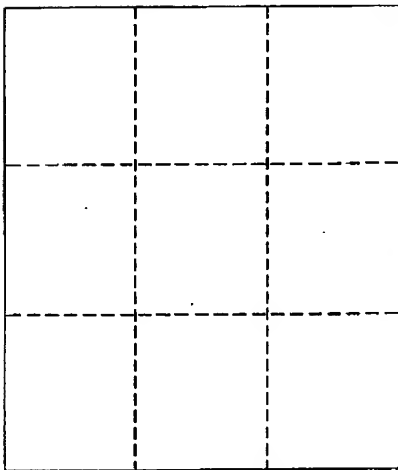
【図9B】



【図2】

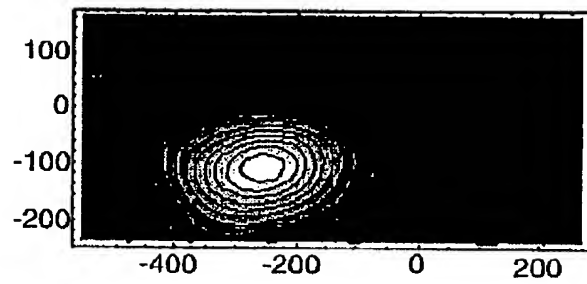


【図6】



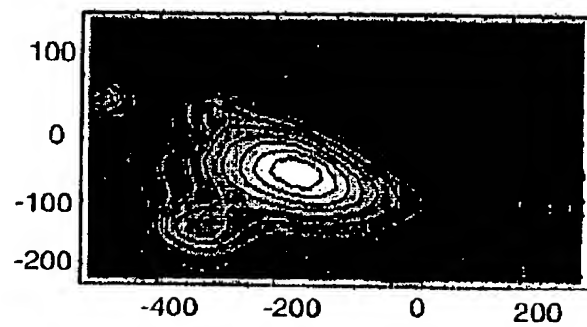
【図8A】

肌STヒストグラム

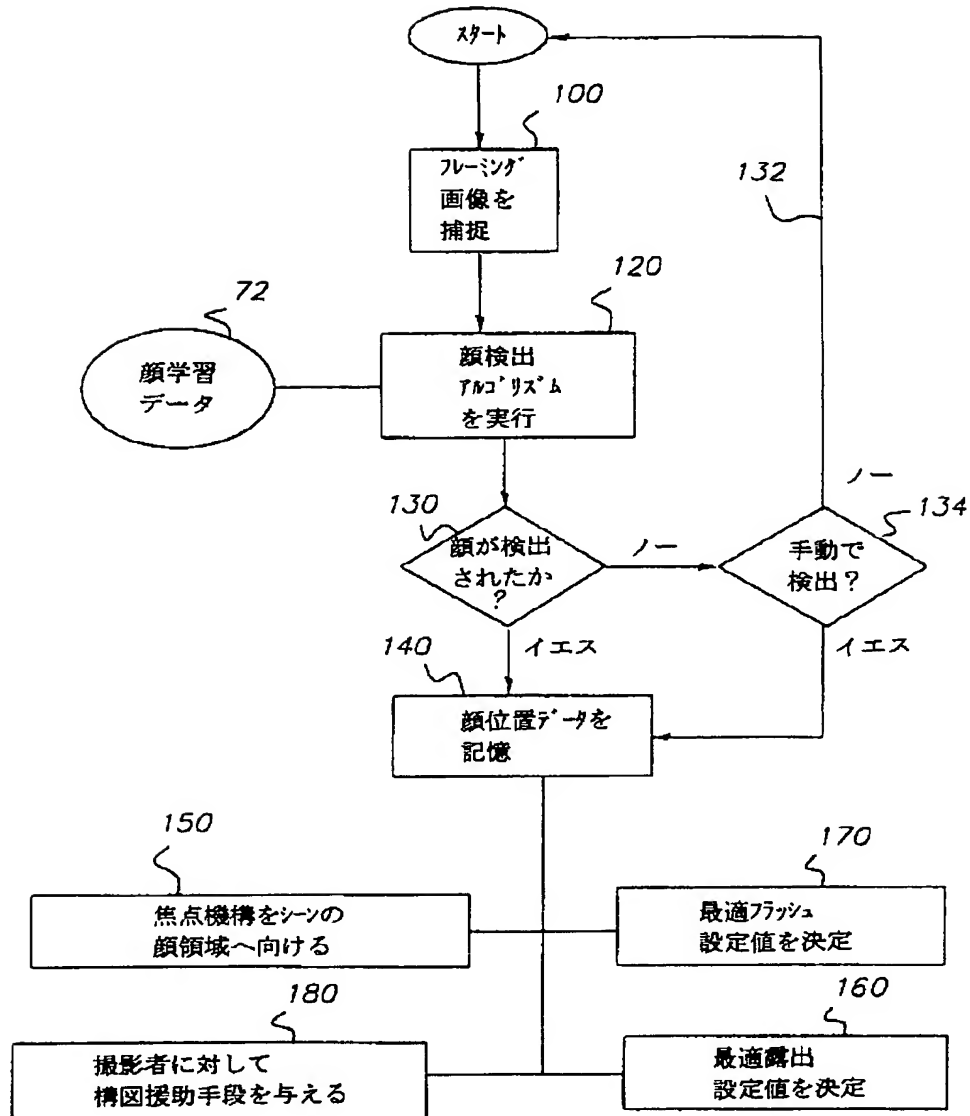


【図9A】

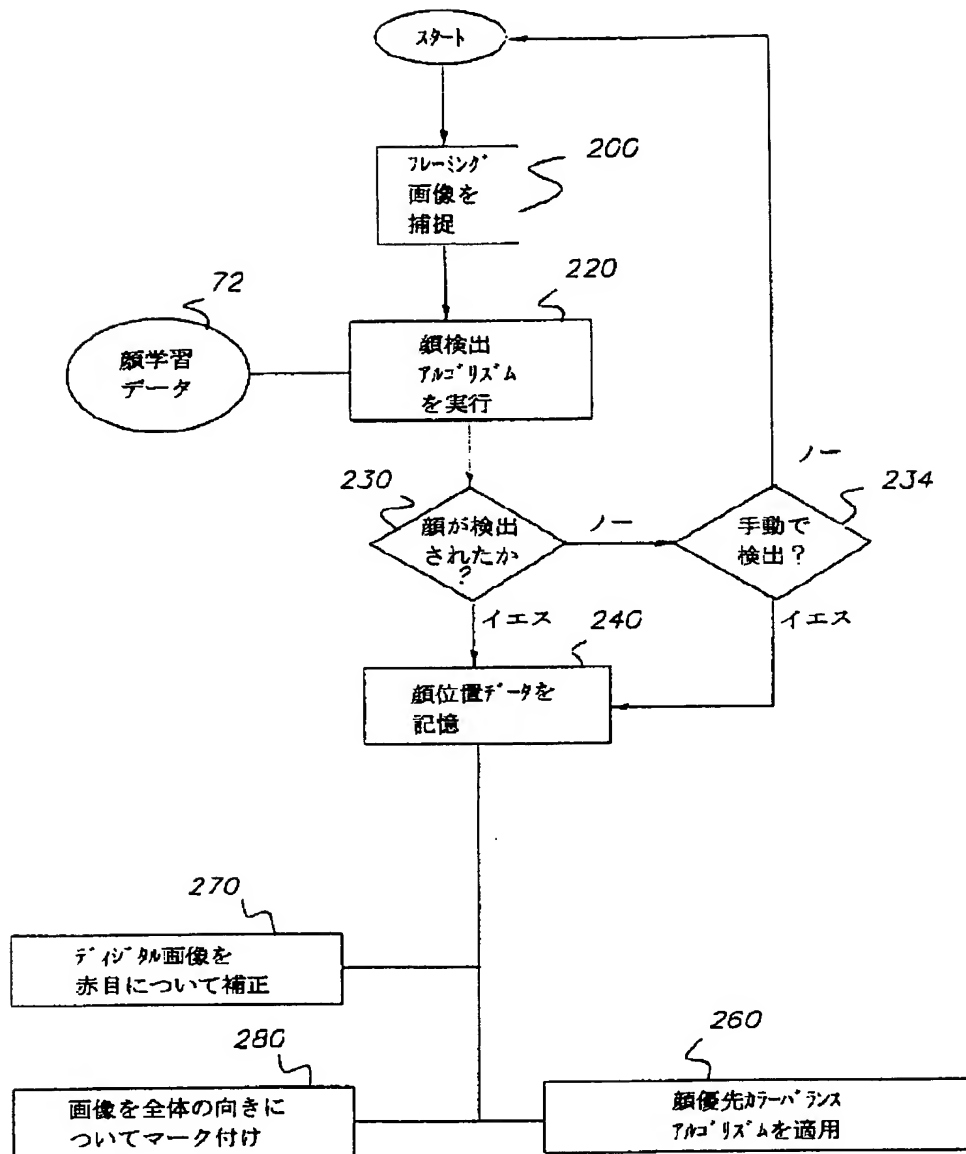
毛髪STヒストグラム



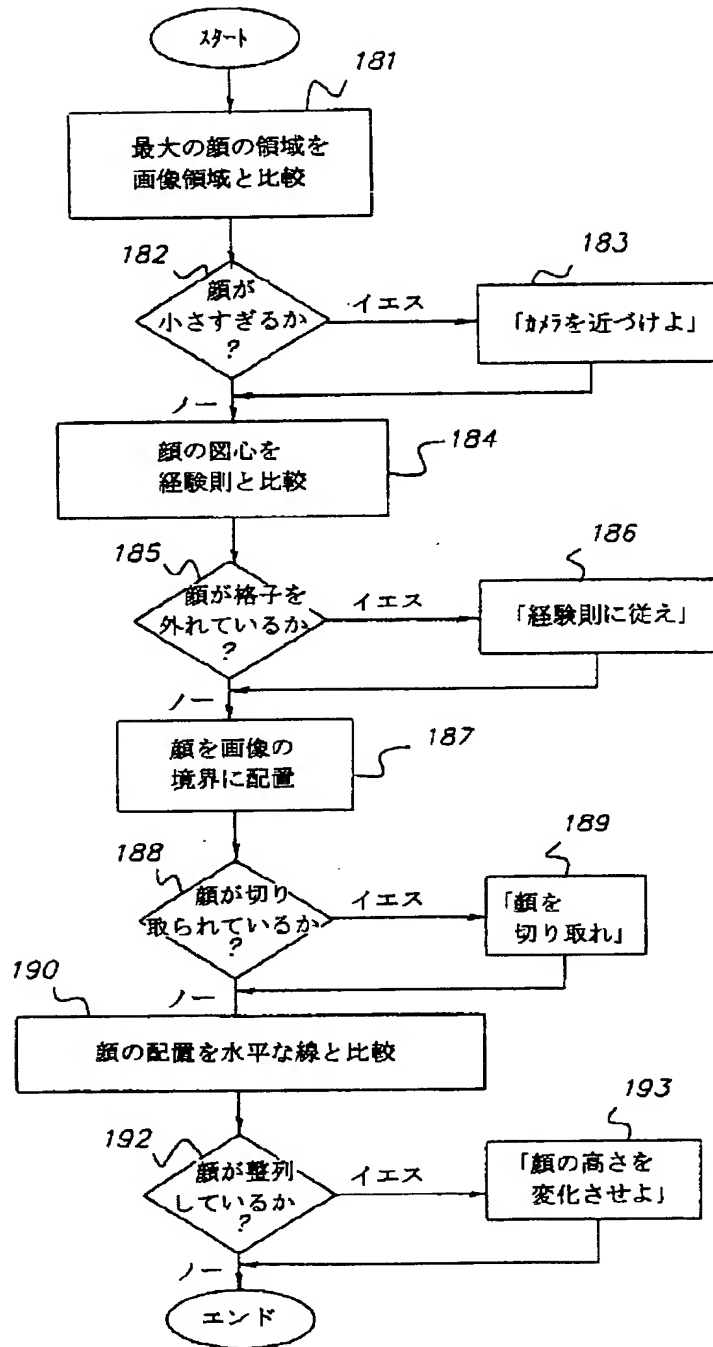
【図3】



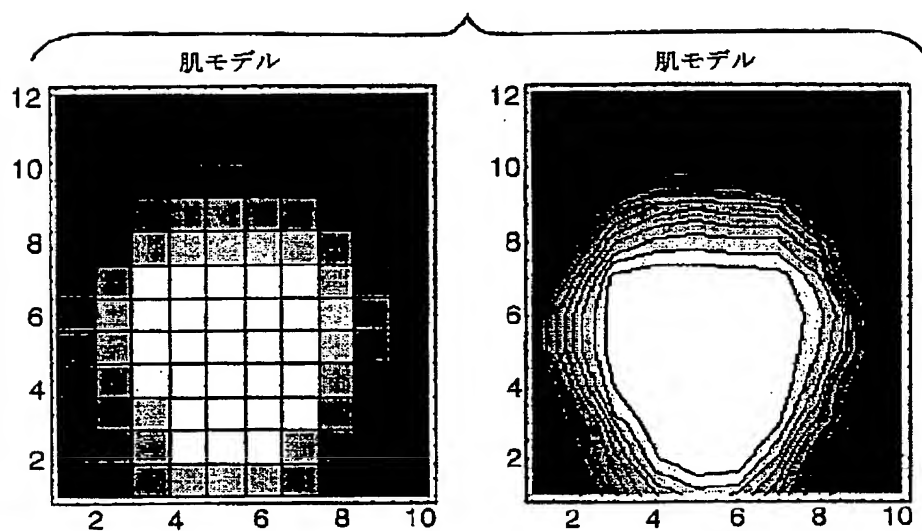
【図4】



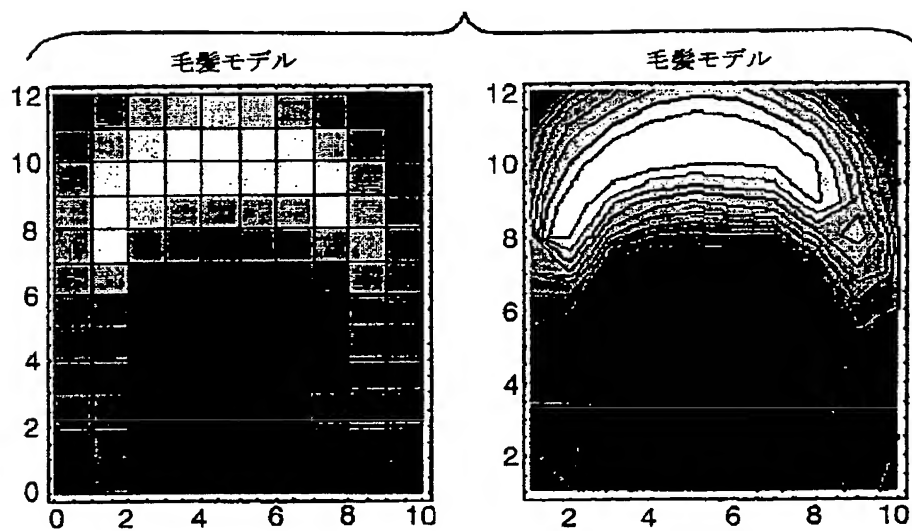
【図5】



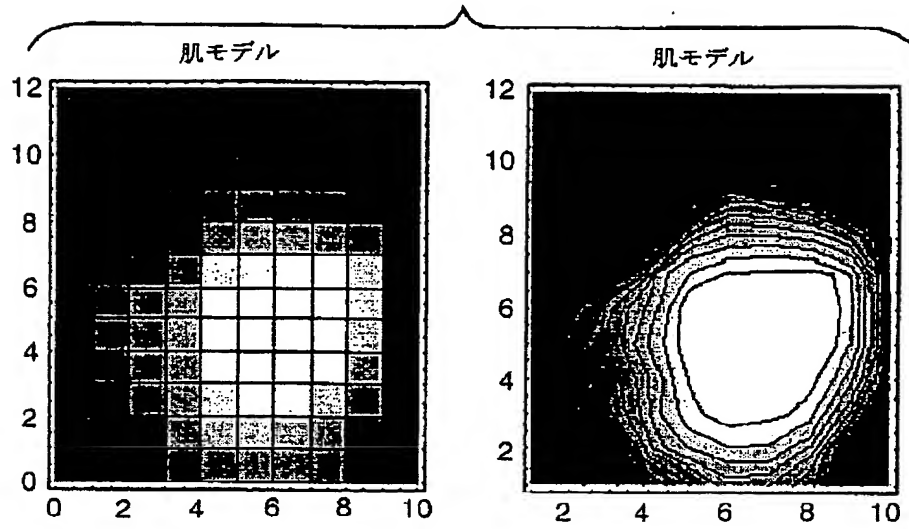
【図7A】



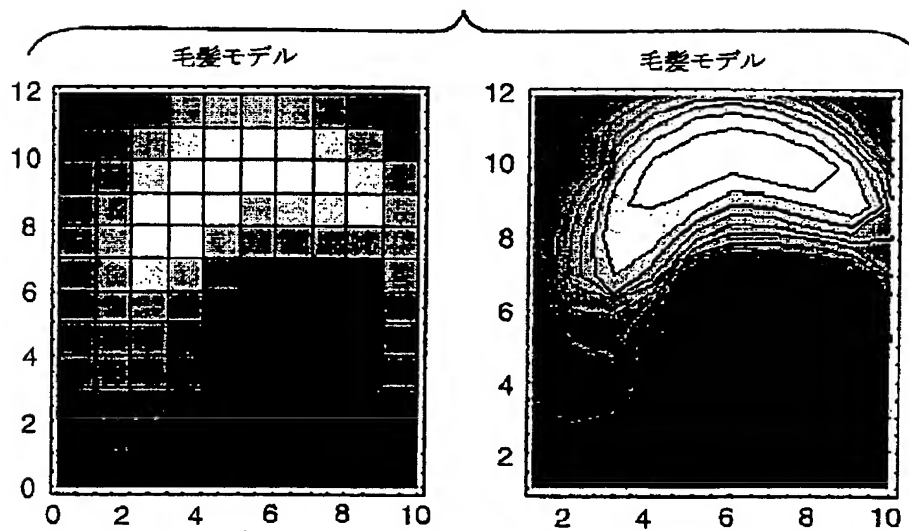
【図7B】



【図7C】



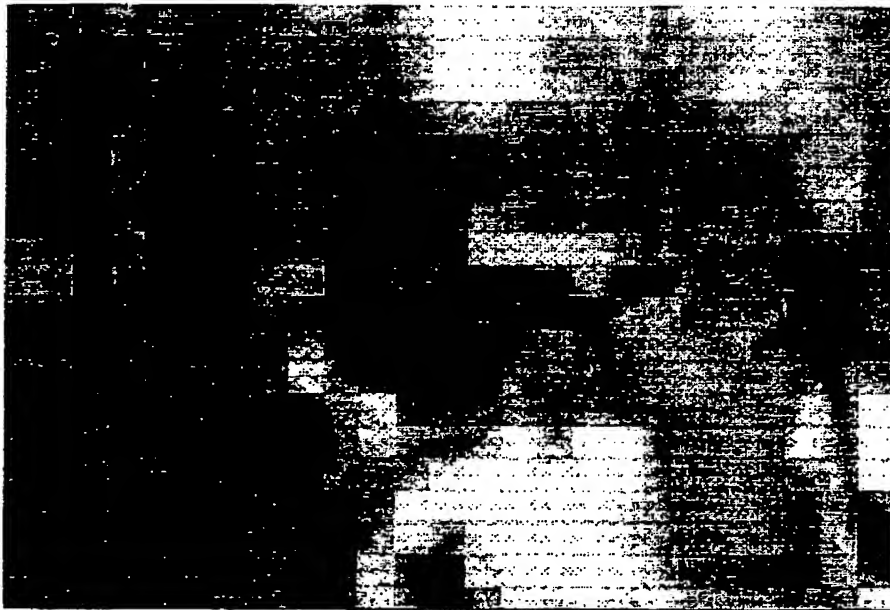
【図7D】



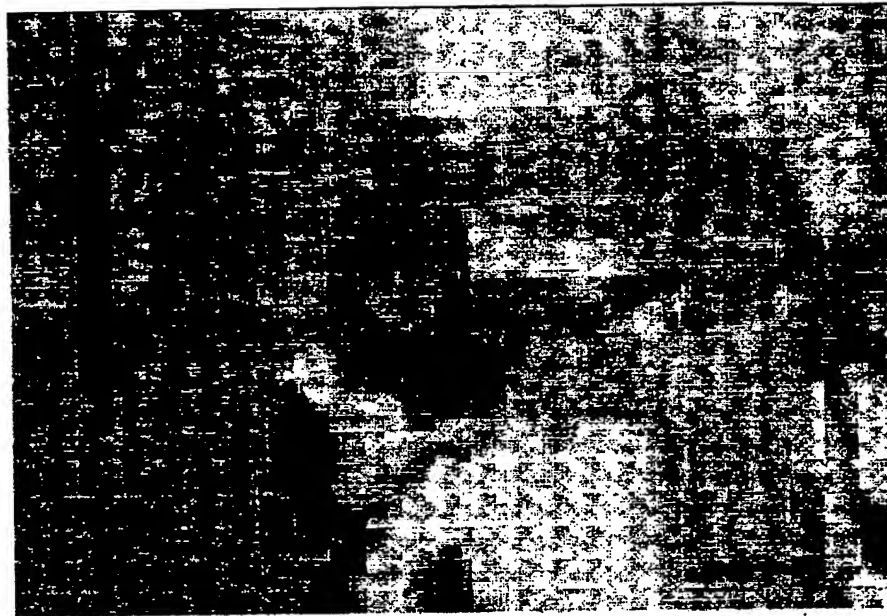
【図10A】



【図10B】



【図10C】



フロントページの続き

(51)Int.Cl.⁷

G06T 7/00

H04N 5/907

// H04N 101:00

識別記号

300

FI

G06T 7/00

H04N 5/907

101:00

サーチワード(参考)

300B

B